

Hand Gesture Recognition using Deep Learning

A. Geetha Devi, M. Aparna, N. Mounika, U. Pavan kalian, R. Meghna Nath



Abstract-- Hearing impaired individuals use sign languages to communicate with others within the community. Because of the wide spread use of this language, hard-of-hearing individuals can easily understand it but it is not known by a lot of normal people. In this paper a hand gesture recognition system has been developed to overcome this problem, for those who don't recognize sign language to communicate simply with hard-of-hearing individuals. In this paper a computer vision-based system is designed to detect sign Language. Datasets used in this paper are binary images. These images are given to the convolution neural network (CNN). This model extracts the features of the image and classifies the images, and it recognises the gestures. The gestures used in this paper are of American Sign Language. In real time system the images are converted to binary images using Hue, Saturation, and Value (HSV) colour model. In this model 87.5% of data is used for training and 12.5% of data is used for testing and the accuracy obtained with this model is 97%.

Key Words: Sign language, HSV colour model, Convolution Neural Networks.

I. INTRODUCTION

Sign language could be a language that gives communication and permits people with hearing or speech impairments to communicate with one another or with other people within the community. In sign language every gesture contains a specific meaning. Therefore, so complicated meanings will be justified by the assistance of combination of varied basic components. There are special rules and grammar's for expressing sign language effectively. Essentially there are two main gesture recognition approaches, image-based and device based. In those two approaches vision-based approach is mostly used, since there is no need of using any sensors or gloves to detect gestures.

Gesture recognition is gaining importance in several applications, areas like human interface, communication, multimedia system and security. Generally, Sign recognition is said as image understanding. It has two steps sign detection and sign recognition. Detection is extracting feature of sure object with regard to sure parameters. Sign recognition is recognizing a definite form that differentiates particular object from the remaining shapes. In this paper we used HSV colour model for skin colour detection and convolution neural network is used for feature extraction.

II. RELATED WORKS

Motion is a non-verbal communication that people use it to communicate feelings and musings. The changed motions of the five fingers and palm may have their physical implications. Hand signal acknowledgment is an entangled framework that is made out of motion demonstration, motion examination and acknowledgment, and Artificial Intelligence (AI). In previous studies on gesture recognition, a sensor is used which senses ultra-wideband impulse signals, reflected from the hand as there is different waveform for every gesture for recognition [1]. Histograms of Oriented Gradients are evaluated for feature extraction where K-Nearest Neighbors (KNN) classifier is used for gesture recognition [2]. Convolution neural network is used for gesture recognition in real time [3]. Data augmentations like zooming, shearing, rotation are applied on datasets before giving it to CNN for better accuracy [4]. There are still impediments with respect to its utilization in troublesome live circumstances where hand gestures variety, illumination change or background change are issues. To avoid these issues Gaussian skin colour model is used to subtract the non-skin colour objects from the images of training and testing data and then pass those to the CNN for better results [5]. For better results hand gestures are determined by using Convex hull Algorithm and then the real-time system is designed by using Neural Network model [6]. CNN will give better accuracy than MobileNetV2 with in less time [7]. Hand gesture recognition can be done on the basis of shape features like centroid, fingers status, and thumb in positions of raised or folded fingers of hand [8]. Custom datasets are used in this approach to make it more accurate and robust when compared with the models that trained with the less number of datasets [9]. Fingertip finder algorithm is used for gesture recognition which is a combination of convex hull and K-curvature for the hand gesture recognition [10].

Revised Manuscript Received on March 11, 2020.

* Correspondence Author

A. Geetha Devi*, Associate Professor, Department of ECE, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, India. Email: geetha.agd@gmail.com

Mekala Aparna, Department of ECE, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, India. Email: m.aparna7135@gmail.com

Nagothi Mounika, Department of ECE, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, India. Email: nagothimounika1998@gmail.com

Udayagiri Pavan Kalyan, Department of ECE, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, India. Email: pavan.kalyan0812@gmail.com

Ramapuram Meghna Nath, Department of ECE, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, India. Email: 1497maggi.ramapuram@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

III. METHODOLOGY

As shown in the Fig.1 hand gesture recognition is divided into different steps, which are image acquisition, image pre-processing, feature extraction, classification, and recognition of hand gesture. In this paper, we have used HSV colour model for skin detection and CNN for feature extraction, classification and gesture recognition.

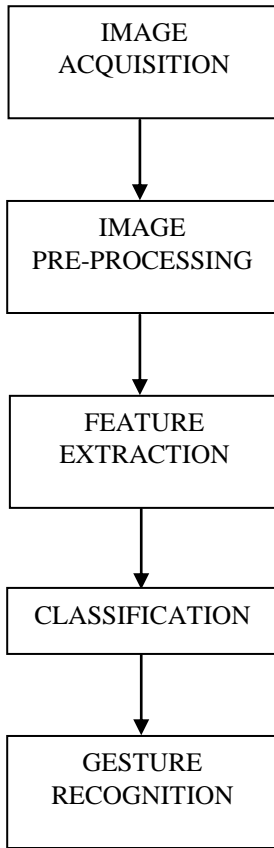


Fig. 1. shows block diagram of hand gesture recognition.

IV. PROPOSED METHODOLOGY

A. Image Acquisition

In this system, the required images are acquired by using webcam. There is a window which will open camera to acquire the images, which shows Region of Interest (ROI) which we use for acquiring hand images.

B. Image Pre-processing

Image Pre-processing is also an important part of hand gesture recognition which is used for resizing and skin colour detection.

- *Resizing*: Acquired images are resized to 64x64 pixels.
- *Skin detection or segmentation*: To indicate existence of humans in an image or video, skin detection is a typical method used in image pre-processing [11]. A real time image is a collection of frames which is captured by web camera. We need to convert RGB images to HSV images, since it is identified with human shading observation. Colour spaces in HSV colour model are divided into three components, H-hue, S-saturation, V-value. To locate the boundaries and hand object, image segmentation is performed. For easy

classification of skin colour and non-skin colour V value gives brightness information. We can specify skin colour boundary in terms of hue and saturation value for HSV features. For accurate boundary of an object and to extract the image properly, adjusting HSV value within range of 0 to 255 is essential. In opencv ranges of hue, saturation, and value are as shown below

H	Hue	0-179
S	Saturation	0-255
V	Vaue	0-255

- *Thresholding*: Thresholding is simple and important part of every work. It is also a method of image segmentation. Binary images are generated from the HSV images which we will get by using thresholding process, In this process pixels of every image is replaced by black pixels if the intensity of those pixels is less than the given threshold values, and pixels with intensity greater than the given threshold value are replaced by white pixels. As intensity is the primary property that pixels of an image can share thresholding images are divided into regions of dark and light.

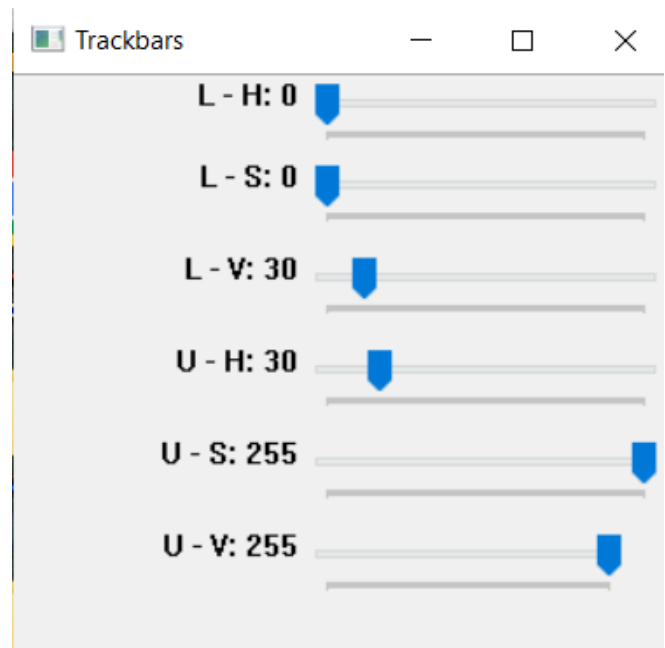


Fig. 2. HSV Track bars to adjust HSV values.

C. Feature extraction

Feature vector of every image which was pre-processed is extracted, by using the designed CNN architecture. Mean-subtracted images are used for feature extraction, which we will get by subtraction of mean-images from the training or testing images [12].

D. Classification

Image processing and morphological process are the steps involved in classifier structure of a general recognition system. With the deep learning techniques, it is possible to classify the images by using less number of image processing and morphological processing units.

In this paper CNN, Tensorflow, and Keras are used for the classification of images. GPUs will help to speed up the process of training and computation, and also make these libraries to work efficiently. CNN based classification is popular in the areas like recognition and image classification. For the activation purpose ReLu function is used which provide good quality being a fast converging tool [6]. Convolution Neural Network which we have used consists of input layer, three 2D Convolution layers, three MaxPooling layers, Flatten, 2 Dense layers, Dropout layer. as shown in fig.3

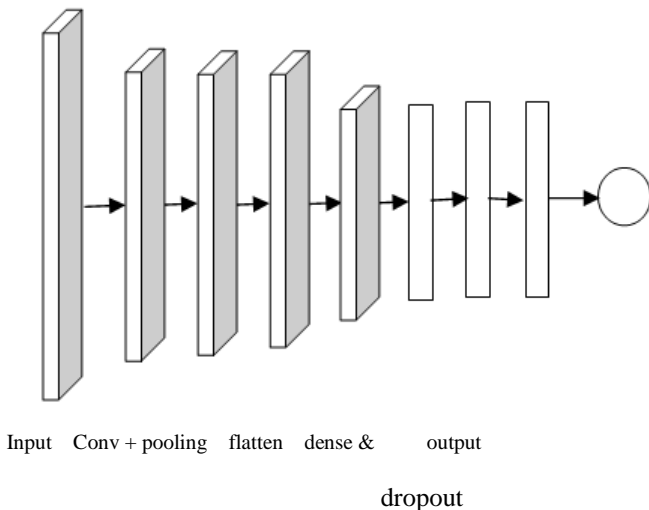


Fig. 3. Architecture of Convolution Neural Network.

E. Gesture recognition

After training process the weights, network model are loaded to the real-time algorithm which has two steps to do one among those steps is extracting hand gestures and the other is classifying those images with CNN. If the image is similar to one of those then the choice is going to be made with the results of those steps. Then the output will be displayed on the screen of the system.

V. RESULT AND DISCUSSIONS

Dataset consists of 52,000 binary images of ASL hand gestures, out of which 87.5% are used for training and 12.5% are used for testing each has 26 classes of 26 alphabets of ASL hand gestures. Images are resized to 64 x 64 pixels, and loaded into program as arrays. Fitting process is started by creating sequential network model, It run through training data with 25 epochs and batch size of 32. Number of images that can be loaded in each iteration is known as batch size where, epochs represents number of cycles that neural network is loaded with images for training. Dropout layer is used to prevent a model from overfitting. Softmax activation function is used in this work, which determines probability of multiple classes at once as it assumes that every class has only one member.

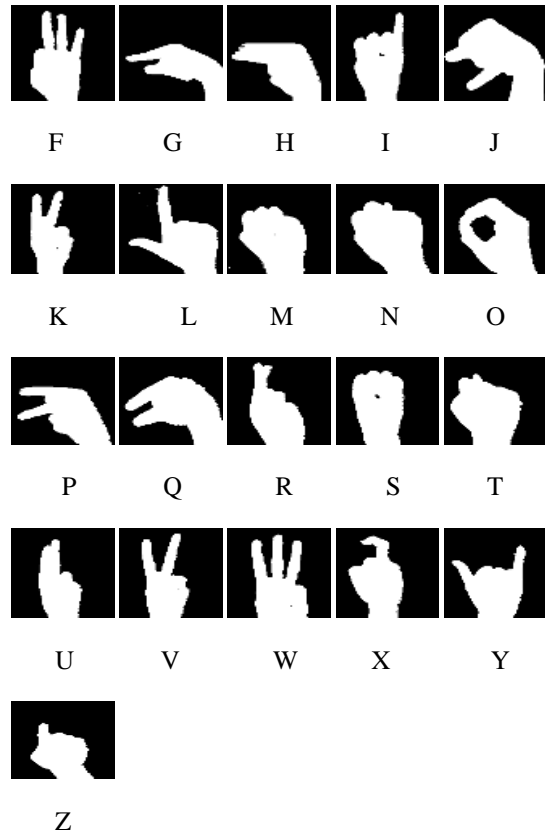
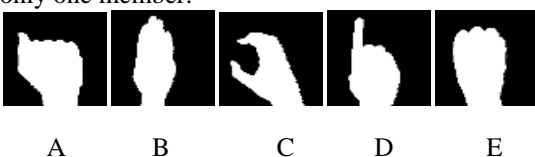


Fig. 4. Binary datasets used for training
Hand gestures we used in this system are of American Sign Language as shown in Fig.4. then the system is converted to real system after the completion of training process.

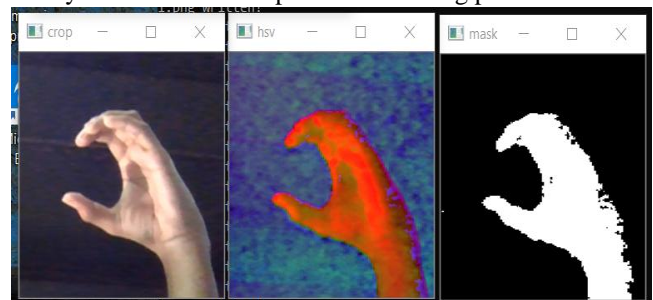


Fig. 5. cropped image, HSV image, binary image
In the real time system gesture image is acquired as cropped image and is converted to HSV image for binarizing the same image. Then the binary image will be tested with the trained model and the output will be displayed as shown in Fig.6



Fig. 6 Output of the hand gesture 'C'

The accuracy and loss of the system while training and testing are as shown in Fig.7 and Fig.8 respectively. The accuracy of the training data which is acc in keras is 97% and the accuracy of the testing data which val_acc is 99%.

VI. CONCLUSION

In this work we developed a real-time system that detects hand gestures and recognize the given gesture. It is to communicate with the hearing-impaired individuals easily. Further, the work can be extended for dynamic hand gesture recognition.

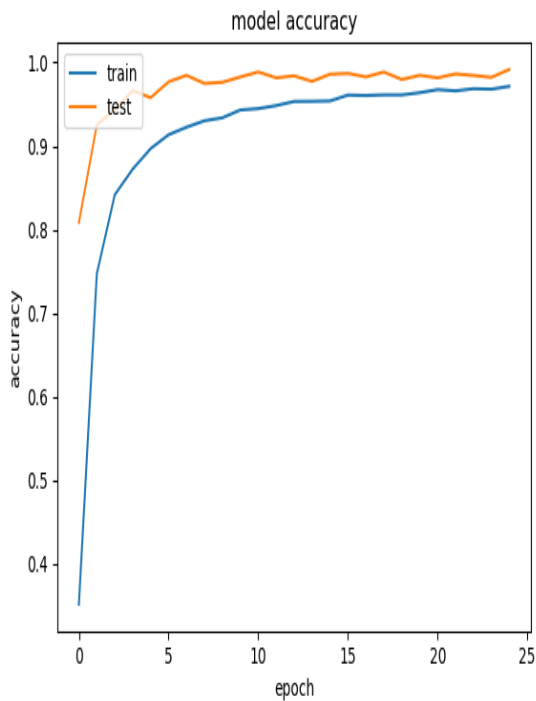


Fig. 7 Accuracy while training and testing

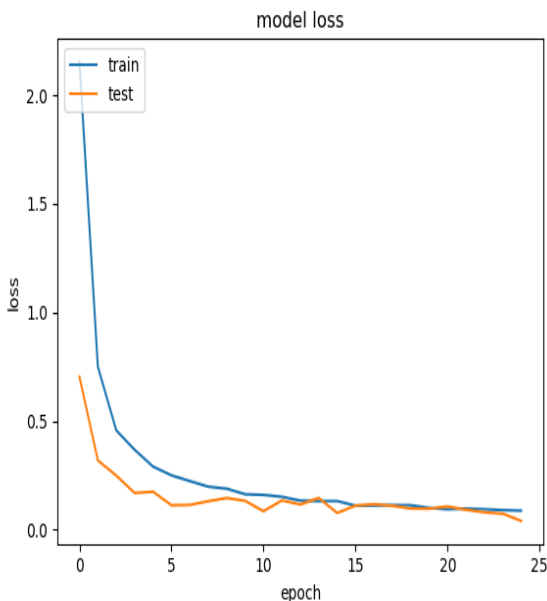


Fig. 8 Model loss while training and testing

REFERENCES

- Seo Yul Kim, Hong Gul Han, Jin Woo Kim, Sanghoon Lee, Tae Wook Kim. "A Hand Gesture Recognition Sensor Using Reflected Impulses".
- [Online]. Available: <https://doi.org/10.1109/JSEN.2017.2679220>
- Daleesha M Viswanathan, Sumam Mary Indicula. "Recognition of Hand Gestures of English Alphabets using HOG Method". [Online]. Available: <https://doi.org/10.1109/ICDSE.2014.6974641>
- Felix Zhan. "Hand Gesture Recognition with Convolution Neural Network". [Online]. Available: <https://doi.org/10.1109/IRI.2019.00054>
- Md. Zahirul Islam, Mohammad Shahadat Hossain, Raihan ul Islam, Karl Andersson. "Static Hand Gesture Recognition using Convolution Neural Network with Data Augmentation". [Online]. Available: <https://doi.org/10.1109/ICIEV.2019.8858563>
- Mengmeng Han, Jiajun Chen, Ling Li, Yuchun Chang. "Visual Hand Gesture Recognition with Convolution Neural Network". [Online]. Available: <https://doi.org/10.1109/SNPD.2016.7515915>
- Murat Taskiran, Mehmet Killioglu, Nihan Kahraman. "A Real-time System for Recognition of American Sign Language by using Deep Learning". [Online]. Available: <https://doi.org/10.1109/TSP.2018.8441304>
- Khalil Bousbai, Mostefa Merah. "A Comparative Study of Hand Gestures Recognition Based on MobileNetV2 and ConvNet Models". [Online]. Available: <https://doi.org/10.1109/ISPA48434.2019.8966918>
- Ashish S. Nikam, Aarti G. Ambekar. "Sign Language Recognition using image-based hand gesture recognition techniques". [Online]. Available: <https://doi.org/10.1109/GET.2016.7916786>
- Manuel Eugenio Morocho Cayamcela, Wansu Lim. "Fine-tuning a pre-trained Convolutional Neural Network Model to translate American Sign Language in Real-time". [Online] Available: <https://doi.org/10.1109/ICCNC.2019.8685536>
- Md. Mohiminul Islam, Sarah Siddiqua, Jawata Afnan. "Real Time Hand Gesture Recognition Using Different Algorithms Based on American Sign Language". [Online] Available: <https://doi.org/10.1109/ICIVPR.2017.7890854>
- A. Sharmila Konwar, B. Sagarika Borah, C. T. Tuithung. "An American Sign Language Detection System using HSV Colour Model and Edge Detection". [Online]. Available: <https://doi.org/10.1109/ICCSP.2014.6949942>
- Byeongkeun Kang, Subarna Tripathi, Troung Q. Nguyen. "Real-time Sign Language Fingerspelling Recognition using Convolutional Neural Networks from Depth map". [Online]. Available: <https://doi.org/10.1109/ACPR.2015.7486481>

AUTHORS PROFILE



Dr. A. Geetha Devi working as Associate Professor in the department of ECE, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada. She has published 14 research papers in various national and international journals and presented 4 papers in national and 9 papers international conferences in the area of image and video processing.



Mekala Aparna, is pursuing Bachelor of Technology in the department of ECE, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada. She is member of IETE student forum and a certified Labview Associate Developer in National instruments.



Nagothi Mounika, is pursuing Bachelor of Technology in the department of ECE, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada. She is member of IETE student forum.



Udayagiri Pavan Kalyan, is pursuing Bachelor of Technology in the department of ECE, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada. He is member of IETE student forum.



Ramapuram Meghna Nath is pursuing Bachelor of Technology in the department of ECE, Prasad V. Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada. She is member of IETE student forum.